mented separately or in combination with one or more other elements. For example, the preprocessor **110** may be implemented as a separate normalization apparatus **200** as will be described below with reference to FIG. **2**.

[0046] When a speech signal to be recognized is input, the preprocessor **110** performs a pre-processing operation for speech recognition.

[0047] For example, when the input speech signal is an analog signal, the preprocessor converts the analog signal into a digital signal, and divides the digital signal into a plurality of speech frames. In one example, a speech signal corresponding to one second is divided into 100 frames. Using data of the divided speech frames, the preprocessor **110** extracts windows to be input to an acoustic model. The preprocessor **110** does not wait until the whole speech signal is input, but extracts a window each time a number of frames corresponding to a window size are input while the speech signal is being input.

[0048] A window is a unit of frame data to be input to an acoustic model, and a window size is a number of frames to be included in a window, and may be set to a value less than the total number of frames of the speech signal to be recognized. The window size may be set and adjusted based on an application field of speech recognition, the computing performance of an apparatus, and other factors. For example, when the application field of speech recognition is translation or interpretation, accuracy may be more important than speed, and the window size may be set to be relatively large. In contrast, when the application field of speech recognition is to process commands embedded in an apparatus such as a TV or other device, the number of commands related to the apparatus is limited, and it may be more important to rapidly process a command. Therefore, the window size may be set to be relatively small.

[0049] When windows for inputting frame data to an acoustic model are extracted in this way, the preprocessor **110** performs normalization in units of the extracted windows. In other words, every time a window is extracted, the preprocessor **110** normalizes the frames belonging to the extracted window.

[0050] The acoustic score calculator **120** calculates acoustic scores in units of the extracted windows using an acoustic model. The acoustic model may be any one of various deep neural network (DNN) acoustic models, for example, a bidirectional recurrent deep neural network (BRDNN) acoustic model. When a window is extracted and frames belonging to the window are normalized by the preprocessor **110**, the acoustic score calculator **120** calculates acoustic scores by inputting the normalized frames into the acoustic model. The acoustic scores may be pronunciation-specific probabilities. However, units for calculating scores are not limited to pronunciations, but may be any one of various language recognition units including, for example, phonemes, morphemes, words, phrases, and sentences.

[0051] The preprocessor **110** may add padding frames to both sides of extracted windows so that the acoustic score calculator **120** may calculate acoustic scores more accurately. It is possible to set in advance whether or not to add padding frames and the number of padding frames to be added. When extracting a current window and adding padding frames, the preprocessor **110** assumes that the added padding frames also belong to the window, and normalizes the frames of the window and the added padding frames together.

[0052] When set to add padding frames, the preprocessor **110** may add one or more left frames adjacent to the first frame of the current window and one or more right frames adjacent to the last frame of the current window. No padding frames are added to the left of the first window and the right of the last window. For example, if two padding frames are to be added, and the current window is a first window including frames **1**, **2**, **3**, and **4**, no frames are added to the left of the current window, and frames **5** and **6** are added to the right of the current window. If the current window is a second window including frames **5**, **6**, **7**, and **8**, frames **3** and **4** are added to the left of the current window, and frames **9** and **10** are added to the right of the current window. If the current window is a last window including frames **13**, **14**, **15**, and **16**, frames **11** and **12** are added to the left of the current window, and no frames are added to the right of the current window.

[0053] When padding frames are added to windows, the acoustic score calculator **120** calculates acoustic scores by inputting together frames of the windows and the added padding frames to the DNN acoustic model. The DNN acoustic model may be a BRDNN acoustic model. Based on an acoustic score calculated for a padding frame added to a preceding window, the acoustic score calculator **120** may calculate an acoustic score for a frame of the current window overlapping the padding frame. In this case, the acoustic score calculator **120** may use an average of acoustic scores calculated for the overlapping frame of the current window and the padding frame, or an average of values obtained by weighting the acoustic scores as the acoustic score for the overlapping frame of the current window.

[0054] For example, assume that the frames of the preceding window are frames **9**, **10**, **11**, and **12**, that frames **7** and **8** were added to the left of frame **9** as padding frames of the preceding window, and that frames **13** and **14** were added to the right of frame **12** as padding frames of the preceding window. Assume that the frames of the current window are frames **13**, **14**, **15**, and **16**. Thus, frame **13** of the current window overlaps padding frame **13** of the preceding window. The acoustic score calculator **120** may use an average of the acoustic scores calculated for frame **13** of the current window and padding frame **13** of the previous window according to the equation $Y=(a+b)/2$, where Y is a recalculated acoustic score of frame **13** of the current window, a is the acoustic score of padding frame **13** of the preceding window, and b is the acoustic score of frame **13** of the current window. Alternatively, the acoustic score calculator **120** may use an average of values obtained by weighting the acoustic scores calculated for frame **13** of the current window and padding frame **13** of the previous window according to the equation $Y=((1-\alpha)a+\alpha b)/2$, where Y is a recalculated acoustic score of frame **13** of the current window, α is a weight value ($0 \leq \alpha \leq 1$), a is the acoustic score of padding frame **13** of the preceding window, and b is the acoustic score of frame **13** of the current window.

[0055] The language score calculator **130** outputs information on words, sentences, and other parts of speech, for example, language scores regarding the frequency of word or sentence use, using a language model and the acoustic scores calculated by the acoustic score calculator **120**. The language model may be an n-gram language model or a neural network language model.

[0056] The interpreter **140** interprets acoustic scores and language scores output from the acoustic score calculator